DOCUMENT RESUME

ED 138 625 TM 006 248

AUTHOR Morra, Linda G.
TITLE Development of a Simulation in Educational Evaluation.
PUB DATE [Apr 77]
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Educational Programs; *Evaluation; Evaluation Methods; *Graduate Study; Higher Education; Models; Programed Instruction; Program Effectiveness; Program Evaluation; School Personnel; *Simulation
IDENTIFIERS Discrepancy Evaluation Model; *Simulation in Educational Evaluation

ABSTRACT

The development of A Simulation in Educational Evaluation (SIMEDEVAL), a simulation designed to increase the evaluation capability of school personnel is described. The Discrepancy Evaluation Model served as the simulation model and the framework for field test procedure. Instruments were constructed and validated for use in the field test with graduate education students. Results indicate that SIMEDEVAL increased students' knowledge of evaluation methodology and willingness to engage in evaluation, and demonstrated their ability to apply evaluation techniques in a structured setting. It is concluded that SIMEDEVAL warrants further development, and that simulations can be subjected to evaluation procedures. (Author/RC)

Development of a Simulation in Educational Evaluation

Linda G. Morra, Ed.D.

State Program Studies Branch

Division of Innovation and Development

400 Maryland Avenue, S.W.

Washington, D.C. 20202

This paper was written and delivered by Linda G. Morra in her private capacity. No official support or endorsement by the U.S. Office of Education is intended or should be inferred.

The study was conducted while the author was a graduate student at the University of Virginia. Further information concerning the study is available through the author.

Introduction

This paper provides an overview of the development of a simulation designed to increase the internal evaluation capability of elementary and secondary education personnel at a relatively low cost in terms of both time and money. Developmental stages include construction of the simulation, development of an evaluation design for the field test of the simulation, instrument construction, and the field test itself.

Construction of the Simulation

A Simulation in Educational Evaluation (SIMEDEVAL) was constructed according to guidelines established by Twelker (1969) and McGuire et. al. (1972). These guidelines may be summarized as a series of steps: (1) define the instructional problem, (2) describe the operational educational system, (3) relate the operational system to the problem, (4) specify objectives in behavioral terms, (5) generate criterion methods, (6) determine the appropriateness of simulation as a method, (7) determine the type of simulation required, (8) determine specifications for simulation, (9) develop the simulation prototype, and (10) field test the simulation prototype. While each step in the construction of SIMEDEVAL will not be detailed, some attention is given to the selection of simulation as the instructional method and the determination of the type of simulation required.

From review of the literature on simulations, four conclusions can be drawn:

--Simulations are interesting to participants (Cruickshank, 1972; Abt, 1970; Taylor & Walford, 1972; Cherryholmes, 1966).

--Simulations can influence participant attitudes and values in a given direction (Heinkel, 1970; Fuller, 1973; Stadsklev, 1973; Coleman, et. al., 1973).

--Simulations are neither more or less effective than traditional methods in teaching factual information or conceptual knowledge (Cherryholmes, 1966; Heinkel, 1970; Fuller, 1973; Stadsklev, 1973; Coleman, et. al., 1973).

--Simulations are most effective when they are used for training purposes, and as a result of the simulation, the participant is expected to perform a specific task in a specific manner (Elder, 1973; Coleman, et. al., 1973).

Thus, simulation is most appropriate as a method of instruction when objectives are to influence participant values and attitudes in a given direction, and also to train participants to perform a specific task in a specific manner. Given the author's objectives of producing a training vehicle which would serve to increase participant's knowledge of evaluation methodology, to broaden their concepts of evaluation, to increase willingness to engage in formal evaluation, and to provide application of specific evaluation techniques, simulation was determined to be an appropriate method of instruction.

The selected format of the simulation can be classified as media-ascendent according to a system developed by Twelker (1969). As a media-ascendent simulation, SIMEDEVAL lacks the common characteristics of interpersonal-ascendent simulations such as emphasis on decision-making, competition, and player interaction. Instead as a media-ascendent simulation, SIMEDEVAL is characterized by the instructional burden carried by media, in this case by programmed instruction. Major advantages of this format (Twelker, 1969) are that the simulation can be used where learners possess low entry skills and limited response repertoires that would prevent an interpersonal-ascendent simulation from well, minimal participation is required by an instructor, learning objectives are generally congruent with standard course objectives, and a relatively large amount of control is possible. Given the intended audience and the objectives of the simulation as well as the desire for a step-by-step incremental training progression, the media-ascendent simulation was selected as most appropriate.

A Simulation in Educational Evaluation is based on the Discrepancy Evaluation Evaluation Model developed by Provus (1971). The Discrepancy Evaluation Model, or the "DEM", includes (1) provision for continuous feedback, (2) procedures for program planning, (3) procedures for examining program implementation, (4) procedures for monitoring program processes, and (5) procedures for program assessment. In the simulation the concept of the team evaluation presented in the DEM is modified so that it describes staff members internal to an educational program who are involved in its evaluation. The simulation also introduces the concept of the participant evaluator, a staff member who is involved in the daily operation of the program and also has responsibility for its evaluation.

Within the simulation the participant plays the role of a public school staff member who needs to develop evaluation skills in order to take responsibility for the evaluation of a Title III program for the gifted. He is assisted in his endeavors by an evaluation consultant who, utilizing a learning-by-doing method, leads the staff member through major steps in the discrepancy evaluation process. The staff member learns to produce a program design (a method of program description) and an evaluation plan for an educational program.

Participants in SIMEDEVAL work through 10 tasks in sequence beginning with a general overview of discrepancy evaluation and exposure to basic terms and concepts and concluding with application of the techniques of program design and evaluation planning. Each of the tasks uses a paper and pencil format and is divided into several segments: a scenario for the task, instructions for the task, materials, response forms, standard response forms against which the participant can compare his/her performance, and a feedback form. A total of ten hours is estimated for simulation completion.

The simulation is designed to serve as a basis of a course or workshop with students working individually or in groups with the supervision of an

instructor or trainer. While SIMEDEVAL has potential use for self-instruction, its recommended use at this time is with supervision.

Evaluation Design

In addition to serving as the framework for the content of the simulation, the Discrepancy Evaluation Model (DEM) also served as the framework for the evaluation of SIMEDEVAL in the field test situation. A Program Design, based on the DEM, was constructed to guide the administration of the simulation, and an Evaluation Plan, based on the same model, was constructed to guide data collection and analysis.

The program design for SIMEDEVAL consisted of two levels. Level I presented an overview of the expected inputs, processes, and outputs; that is, it described the resources or preconditions necessary for field test of the simulation, the major activities or tasks within the simulation, and the expected terminal or final outcomes for the simulation. Level II detailed the inputs processes, and enabling or interim objectives for each task of the simulation. The program design not only served as a procedural guide for administration of the simulation, but also served as the basis for evaluation by providing the standard against which actual performance could be compared. The evaluation plan for SIMEDEVAL delineated the areas of concern addressed by the field test, the specific evaluation questions asked under each area of concern, and for each evaluation question, the standard against which the performance was to be compared, the source of the information, the instrument to be used to collect the information, proposed data analysis, and date(s) for instrument administration.

In brief, the evaluation design for SIMEDEVAL was intended to provide information as to the initial value or face validity of the simulation, the effectiveness of the simulation within the field test situation, and needed

areas of revision within the simulation. To determine the effectiveness of the simulation within the field test, four hypotheses (also terminal objectives of the simulation) were tested:

--There will be a significant difference in participants' scores on a pre- and post-test measure of evaluation concepts.

--There will be a significant difference in participants' scores on a pre- and post-test measure of willingness to use formal evaluation procedures.

--There will be a significant difference in participants' scores on a pre- and post-test measure of knowledge of evaluation methodology.

--Participants can demonstrate ability to apply evaluation techniques by producing a program design and evaluation plan which meet pre-determined criteria.

Instrumentation

Instruments to be developed for the field test of SIMEDEVAL included feedback forms, time sheets, biographic information forms, as well as instruments for determination of the effectiveness of the simulation. An Evaluation Test was constructed to determine each simulation participant's knowledge of evaluation methodology, and an Evaluation Attitude Questionnaire was constructed to determine both participants' concept of evaluation and willingness to engage in formal evaluation. An Evaluation Rating Scale was also developed to determine initial value of the simulation. Criteria for determining the adequacy of each participant's program design and evaluation plan were included within the simulation, although additional scoring guidelines for use of outside raters were developed.

The Evaluation Test, used as the measure of knowledge of evaluation methodology, consisted of 16 multiple choice items. The items were derived from analysis of texts by Worthen and Sanders (1973) and the PDK National Study Committee on Evaluation (1971), and were based on the identification of major steps necessary in conducting formative and summative educational

evaluation. Information concerning the reliability and validity of the instrument was obtained from 28 gradu te students in education at the University of Virginia. A reliability coefficient of .79 was obtained for the Evaluation Test. In addition, 8 of the test's items were found to discriminate at a .05 level of significance between graduate students with and without a major in educational evaluation.

The Evaluation Attitude Questionnaire, used as the measure of concept of evaluation and indicator of willingness to engage in formal evaluation, consisted of two scales. The first, or "concept" scale consisted of 16 items derived from Wolf's (1972, 1973) work with public school teachers. Wolf identified 9 variables which discriminate between individuals holding a broad or narrow concept of evaluation. For example, a broad concept of evaluation is held by teachers who consider the judgements and opinions of students important in evaluation of programs. As adapted for the Evaluation Attitude Questionnaire, participants were to indicate their agreement/disagreement on a 5-point scale with the importance of student opinion in program evaluation. While a reliability coefficient of .82 was obtained for the concept scale, the scale failed to discriminate between evaluation and non-evaluation graduate students.

The second, or "willingness" scale consisted of 2 items. Using a 5-point rating system, participants were to indicate their willingness to engage in formal program evaluation. A reliability coefficient of .82 was obtained. In addition, these items did discriminate between evaluation and non-evaluation graduate students.

As a measure of the initial value of the simulation, a Simulation Rating Scale was developed. The scale was developed from an analysis of the literature which resulted in the identification of 10 criteria for determining the initial value of the simulation: instructional value, validity, comprehensiveness,

verisimilitude, symmetry, synchronization, manageability, ease of administration, immediacy of feedback, and presentation of learning objectives. These criteria were used as the basis for the rating scale.

Field Test of the Simulation

Methods

The field test of the simulation took place at the School of Education, San Jose State College, SIMEDEVAL was offered to graduate level students as a Special Studies one credit, 16 hour course. The course was taught by the author from 1:00 to 5:00 pm daily, June 24 through June 27, 1975. At the beginning of the first session, participants were asked to complete the pre-measures (Biographic Data Questionnaire, Evaluation Attitude Questionnaire, and the Evaluation Test). Evaluation Attitude Questionnaires were precoded to allow for anonymity of response. After an introduction to the simulation by the instructor, participants were allowed to work individually, proceeding at different rates. Neither formal teaching nor discussion sessions were offered by the instructor, although individual or group requests for clarification of a task were met and recorded. None of the simulation material was used for formative purposes during the field test, that is, a student who did not meet the standard on Task One was not given remediation before continuing with Task Two. Upon completion of the simulation, participants were asked to complete post-measures (Simulation Rating Scale, the precoded Evaluation Attitude Questionnaire and the Evaluation Test).

During the week of July 21, 1975, evaulation faculty, staff, and advanced graduate students were asked to examine SIMEDEVAL briefly and rate its initial value.

During the week of July 28, 1975, two doctoral level graduate students in evaluation and the author independently scored the program designs and evaluation plans produced by the simulation participants in Tasks 9 and 10 of the simulation. Each rater was given a set of scoring guidelines. Each Participant's final score was the average of the 3 ratings for each task.

Results

The 10 students participating in the field test of SIMEDEVAL were typically female, between 41 and 50 years of age, and elementary school teachers with many years of teaching experience and graduate level coursework. Over half of the students had been part of writing teams whose task was to develop Early Childhood Education or Individually Guided Education program or program designs.

Three questions were asked of the field test of SIMEDEVAL. (A) What changes are indicated either in the simulation or its administration to improve its effectiveness?, (B) How effective is the simulation within the field-testing situation?, and (C) What is the initial value of the simulation? In response to the first question, the field-testing of the simulation indicated several areas of needed revision. These included revision of several lengthy tasks to bring the simulatin within the 10 hour desired completion time, clarification of instructions and response formats for several of the tasks, and modifications in the author's standards for several of the tasks.

Determination of the effectiveness of the simulation required examination of the 4 hypotheses stated previously. First, as a measure of the effectiveness of the simulation, the hypothesis was stated that students can produce both a program design and an evaluation plan for an educational program which meets a predetermined standard. Two doctoral level graduate evaluation students and the author independently scored the simulation participant's products using pre-established scoring guidelines and criteria. Participants were

assigned the average of the 3 ratings for each task . All 10 participants met the standards established for the products.

As the second indicator of the effectiveness of SIMEDEVAL, the hypothesis was stated that there would be a significant difference in participant's scores on a pre- and post-test measure of knowledge of evaluation methodology. Only those items of the Evaluation Test which were found to discriminate between graduate students in evaluation and non-evaluation graduate students were used in the analysis. Results indicated a difference in the participants' scores on the pre- and post- measure of knowledge of evaluation methodology significant at the .05 level.

As a third indicator of the effectiveness of the simulation, the hypothesis was stated that participants' scores on a pre- and post-test measure of concept of evaluation would differ. While at the .05 level of significance, such a difference was found, the lack of demonstrated validity of the Evaluation Concept Scale precluded acceptance of the hypothesis that simulation participants broaden concepts of evaluation.

Finally, as the fourth indicator of the effectiveness of SIMEDEVAL, the hypothesis was stated that simulation participants' scores on a pre- and post-test measure of willingness to engage in formal evaluation would differ significantly. Results indicated that participants in the simulation did increase in willingness to engage in formal program evaluation at a .05 level of significance.

The third question asked in the field test of SIMEDEVAL concerned the initial value or face validity of the simulation. Simulation participants and evaluation graduate students, staff, and faculty gave SIMEDEVAL high ratings on all 10 criteria presented in the Evaluation Rating Scale. In addition, many of the simulation participants commented that it had been an interesting, valuable, and challenging experience.

Conclusions

The development of SIMEDEVAL offers three major contributions. First, the study indicates that an evaluation model such as the DEM can be used to evaluate simulations. Application of an evaluation model can provide information to improve the simulation itself as well as to assess its effectiveness in a given situation. Second, the work has produced an evaluation knowledge test and a scale which measures willingness to engage in formal evaluation which have demonstrated reliability and validity. Beyond their application to SIMEDEVAL, the instruments may be useful in examining the effectiveness of other evaluation training procedures. Third, SIMEDEVAL looks like a promising vehicle for increasing the internal evaluation capability of elementary and secondary school personnel. While investigation has been limited to one group of students, the simulation was effective in increasing participants' knowledge of evaluation methodology, willingness to engage in formal evaluation, and ability to produce program designs and evaluation plans which encompass both formative and summative evaluation procedures. In addition, SIMEDEVAL has initial value in terms of criteria established for good simulations as rated by both simulation participants and evaluation specialists. In sum, it can be concluded that A Simulation in Educational Evaluation merits further development and testing.

References

Abt, Clark C. Serious Games. New York: Viking P

Cherryholmes, Cleo H. "Some Current Research on Effectiveness of Educational Simulations: Implications for Alternative Strategies." American Behavioral Scientist, X (June, 1966), 4 - 5.

Coleman, James S., Livingston, Samuel A., Fennessey, Gail M., Edwards, Keith J., and Kidder, Steven J. "The Hopkins Games Program: Conclusions from Seven Years of Research." John Hopkins University, Center for Social Organization of Schools, August, 1973.

Elder, Charles D. "Problems in the Structure and Use of Educational Simulation." Sociology of Education, XLVI (February, 1973), 335 - 5.

Fuller, Jack W. "Simulation in a Political Science Classroom." Improving College and University Teaching, XXI (March, 1973), 284 - 5.

Heinkel, Otto A. "Evaluation of Simulation as a Teaching Device." Journal of Experimental Education, XXXVIII (January, 1970), 32 - 6.

McGuire, Christine H., Solomon, Lawrence M., and Bashook, Philip G. Handbook of Written Simulations: Their Constructiona and Analysis. University of Chicago: Center for Educational Development, 1972.

PDK National Study Commission on Evaluation. Educational Evaluation and Decision Making. Bloomington, Indiana: Phi Delta Kapp, Inc., 1971.

Provus, Malcolm. Discrepancy Evaluation. Berkeley, California: McCutchan Pub., 1971.

Stadsklev, Ronald. A Comparative Study of Simulation Gaming and Lecture-Discussion Method. (1973) ERIC, ED 065 405.

Twelker, Paul A. "Designing Simulation Systems." Educational Technology, IX (October, 1969), 64 - 70.

Wolf, Robert L. "The Role of the Teacher in Classroom Evaluation." Unpublished Master's Thesis, Center for Instructional Research and Curriculum Evaluation, University of Illinois, 1972.

Wolf, Robert L. "How Teachers Feel Toward Evaluation." School Evaluation: The Politics and Process. Edited by Ernest R. House. Berkeley, California: McCutchan Pub., 1973.

Worthen, Blaine R., and Sanders, James R. Educational Evaluation: Theory and Practice. Worthington, Ohio: Charles A. Jones, 1973.